AD_____

AWARD NUMBER: DAMD17-00-1-0431

TITLE: Breast Cancer Susceptibility Genes in High Risk Women

PRINCIPAL INVESTIGATOR: Ann S. Hamilton, Ph.D.

CONTRACTING ORGANIZATION: University of Southern California
Los Angeles, California 90033

REPORT DATE: December 2005

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-12-2005	Final	1 Jul 2000 – 30 Nov 2005

4. TITLE AND SUBTITLE

Breast Cancer Susceptibility Genes in High Risk Women

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-00-1-0431

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Ann S. Hamilton, Ph.D.

E-Mail: ahamilt@usc.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Southern California
Los Angeles, California 90033

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A positive family history, present in about 30% of breast cancer cases, has been shown to double a woman's risk of breast cancer. The genetic factors responsible are largely unknown, although the relatively high penetrant genes BRCA1/2 may account for 3%. Genes of lower penetrance may also affect breast cancer risk, and a likely group of such genes are those that regulate the production, intracellular transport, and metabolism of estrogen. Previous studies of these susceptibility genes have not been conducted with women with high familial risk. This study included identical twins with differing genetic risks (i.e. concordant for breast cancer pairs vs. discordant pairs) as well as unaffected controls. DNA samples were obtained from 136 concordant pairs, 152 discordant pairs and 137 controls. DNA has been extracted and stored for conducting additional genetic testing using these samples. A total of 368 single nucleotide polymorphisms (SNPs) have been assayed along 16 genes using the Illumina System. The genes included AIB1, COMT, COX2, CYP17, CYP19, CYP1A1, CYP3A4, ESR1, ESR2, GPR54 ,GSTP1, IGF1, IGFBP3, P160, and PR. The SNPs selected were haplotype tagging SNPs that were selected to cover the variation across the entire length of each of the genes. Genes preliminarily showing the most indication of being involved with breast cancer susceptibility included HSD17B1, CYP1A1, GSTP1, AIB1, P160 and COX2. The project has generated a wealth of data that will require further analysis to understand the significance of these results and has created a valuable resource for additional testing of newly identified SNPs.

15. SUBJECT TERMS
Breast cancer ,Twins, estrogen metabolism, carcinogen metabolism, co-activators, genetics

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	23	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

4) INTRODUCTION

A positive family history, present in about 30% of breast cancer cases, has been shown to double a woman's risk of breast cancer(1), and this is true for postmenopausal as well as the premenopausal cases, among which the autosomal dominant, relatively high penetrant genes BRCA1 and BRCA2 are most prominent(2). It has been hypothesized that susceptibility genes of lower penetrance are more prevalent than among the latter, and a likely group of such genes are those that regulate the production, intracellular transport, and metabolism of estrogen (3), the common factor underlying most known predictors of breast cancer risk (4) (5) (6).

At the time the grant was originally funded, reviews identified several candidate genes (7) (8) (9) that were hypothesized to be related to genetic risk of breast cancer. The original proposal was limited to investigating single nucleotide polymorphisms (SNPs) on 6 genes that were related to estrogen metabolism and carcinogen metabolism. Specifically, in the estrogen metabolism pathway, four genetic polymorphisms were previously described related to the CYP17 gene, the CYP19 gene, the COMT gene, and the HSD17B1 (or also called the EDH17B2) gene. For example, a polymorphism (called A2) on the CYP17 gene was linked to higher endogenous estrogen levels and an earlier age at menarche (10). The same polymorphism was linked to increased risk of aggressive breast cancer, although one attempt to confirm this finding was unsuccessful (11). Genes related to carcinogen metabolism which have been linked to breast cancer risk include GSTM1 and P1 and CYP1A1. As a result of rapidly improving technology, for the same cost, the study was able to expand the original goals and greatly increase the number of genes and SNPs being investigated. Specifically, there are 16 genes and 384 SNPs that were assayed. The genes include AIB1, COMT, COX2, CYP17, CYP19, CYP1A1, CYP3A4, ESR1, ESR2, GPR54 ,GSTP1, IGF1, IGFBP3, P160, and PR. The SNPs selected were essentially haplotype tagging SNPs that were selected to cover the variation across the entire length of each of the genes.

Most of the previous studies of genetic polymorphisms have not been conducted with women known to be at high familial risk of breast cancer, where the prevalence of the polymorphism may be expected to be higher, if it is associated with the development of breast cancer. The identification of families to study these inherited genetic factors is more difficult because of the anticipated lower penetrance of the candidate genes and occurrence of more sporadic cases, especially among older women. The source of the breast cancer cases in this study was the International Twin Study which includes both breast cancer concordant and discordant identical twin pairs. The concordant MZ twin pairs represent families with a very high familial risk of breast cancer, while the MZ discordant twins are likely to represent non-heritable cancer (36). We have obtained DNA from subsets of these pairs as well as from control women without breast cancer (and without a family history of breast cancer) and have tested for multiple genetic polymorphisms to determine if any are differentially associated with cases from twins with a high likelihood of heritable breast cancer (i.e. those from identical concordant pairs) vs. sporadic cases (i.e. cases from discordant pairs) and control women.

This study should provide important clues regarding other genetic factors that may be associated with breast cancer etiology. A previous publication by the P.I. on epidemiological risk factors within the concordant for breast cancer identical twins, has indicated that factors associated with the onset of hormones at puberty may be especially critical (33). The twins are a unique subset of very high risk women who have been identified by the fact that they are identical twins from pairs in which both members have developed breast cancer. If multiple genes are involved (which is likely) this group of cases is an especially important resource to study since the different genes responsible may have been passed down separately from each parent. Thus there may have been no (other) family history of breast cancer for these twins since each gene, by itself, may not have been sufficient to increase risk of disease. Initial work on the project and the CYP17 laboratory work was funded under a grant from the California Breast Cancer Research Project (CA-BCRP).

As indicated above, during the no-cost extension period we increased the number of genetic factors studied using these twins with new high through-put technology that has recently become available at USC. Details are included in the section below under Task 4.

5) BODY

Technical Objectives and Work Accomplished During Grant period

Task 1: To complete follow-up of female identical twin pairs with breast cancer

1. *Continue follow-up begun under CA-BCRP grant*
2. *Hire Programmer, set up tracking database*
3. *Continue to mail follow-up forms with return envelope to last known address of twins. Enter data from responses.*
4. S*ubmit nonrespondent names to National Death Index.*
5. *Submit names of nonrespondent twins not known to be deceased to TRW/ Experian to obtain updated addresses. Resend follow-up forms.*
6. *Continue follow-up by phone calls, internet searches, and contact with relatives.*

It was previously reported that a data file was created from the International Twin Registry that selected all of the identical female twin pairs in which one or both members had been diagnosed with breast cancer. In total there were 1,491 identical pairs in this database and 1,199 of them were initially classified as discordant pairs, 263 as concordant, and 29 of uncertain concordance. A follow-up form was sent to all living members of all of the discordant pairs, and new breast cancers have been reported in the previously healthy twin of 62 of these pairs. Thus as a result of this information, we identified 338 concordant pairs and 1,153 discordant pairs. Follow-up efforts have consisted of mailing 1,883 follow-up forms to living twins in these pairs, and 1,029 have been returned completed. 260 were returned by the post office and 478 were not returned by either the twin or the post office. Tracing efforts were implemented to locate the nonrespondents. Follow-up of all nonrespondents was done using the National Death Index. (This component was funded under the CA-BCRP grant).

Task 2: Identify new breast cancers and obtain medical record documentation and tissue blocks.

1. *When new breast cancer is identified, obtain medical consent form from twin or next of kin, and request records and tissue blocks from hospital*
2. *Follow-up requests with hospitals*

The goal of the study was to obtain genomic DNA from at least one member of 200 of the concordant pairs, from the case in 200 of the discordant pairs, and from 200 control women without a personal or family history of breast cancer. From a previous study, tissue blocks have been obtained from some of the breast cancer pairs (concordant and discordant). As a result of the follow-up effort, we have identified 62 previously discordant pairs in whom the unaffected member has developed breast cancer. Thus the number of concordant and discordant pairs has been adjusted to reflect the current status.

To participate in the study, the eligible participants were sent a letter describing the study along with the informed consent documents. Our study manager then called the twin to go over the informed consent with her over the telephone. Then if she agreed to participate and donate the required tissue to the study, she signed the informed consent form and mailed it back to us. .

The numbers of MZ twins (and controls) in each subset with tissue and signed consent forms is the following:

	Concordant	Discordant	Controls
Number identified *	169	892	--
DOD consent signed and tissue/buccal smear available	**136**	**152**	**137**
(Number of above with buccal smear)	(42, with 13 having both buccal and tissue)	(20, with 8 having both buccal and tissue))	(137)

*after elimination of refusals, and deceased cases with no available tissue. Reasons for refusal included not interested, and too busy as well as the language that the DOD requires us to include in the informed consent regarding 'POTENTIAL FOR COMMERCIAL DEVELOPMENT RELATED TO RESEARCH'.

We completed the study with tissue or buccal smears and signed DOD informed consents for 136 concordant pairs, 152 discordant pairs and 137 controls. Due to difficulty in locating subjects it took more staff time than anticipated to obtain the current numbers.

Task 3: Obtain buccal smears from living member of case pairs when blocks not available

1. *If tissue blocks are no longer available from either member of the case pairs and there is a living twin, send letter to obtain buccal smear.*
2. *Send buccal smear kit and return mailing supplies and postage to these individuals.*

The procedures for obtaining buccal smears have been developed and kits were assembled for this purpose. We used Epicentre Technologies Master Amp Buccal Swab Brush. Two brushes

were sent to the selected cases (and controls) and they were asked to use one for each cheek. Once the swabs were returned to us they were kept frozen until the laboratory analyses were done. We collected buccal smears from 42 concordant pairs, 20 discordant pairs and 137 controls.

Task 3: Identify 200 control women and obtain buccal smear and risk factor questionnaire from each of them

1. *Contact case pairs to obtain listing of unrelated breast cancer free potential control women selected from sisters-in-laws and friends.*
2. *Randomly select a women from this list and mail introductory letter.*
3. *Obtain buccal smear and risk factor questionnaire from each control woman through the mail.*

We developed the protocol for selecting controls and this is worked well. We identified 137 controls and have obtained the buccal smear and short risk factor questionnaire from all of them.

Task 4: Laboratory analysis of DNA from tissue and buccal smears to identify polymorphisms in the specified breast susceptibility candidate genes

1. *Finish CYP-17 analysis at Dr. Dubeau's Laboratory.*
2. *Extract additional DNA as necessary for the additional genetic tests.*
3. *Do additional tests for CYP19, COMT, HSD17B1, GSTM1, GSTP1, and CYP1A1.*
4. *Receive results and enter data into database.*
5. *Store tissue for future genetic studies.*

We had some difficulties in this area have worked to resolve the problems. This caused some delay in completing the genetic analyses. During this time period technological advances have been made in doing genetic assays, and costs per assay have been reduced. These developments provided the opportunity to expand the scope of the genetic analyses that could be done with the available funding.

Haplotype tagging SNPs (htSNPs) were selected to predict the common haplotypes in each gene with a high probability (R^2 =.80), similar to methods described in studies done using the Multiethnic Cohort (37). We included the 6 genes from the original proposal and added 10 more genes to the study based on current research findings and biological plausibility. The additional genes, listed in the table below include the co-activators AIB1 and p160, IGF related genes including IGF-1 and IGFBP-3 which regulates the amount of IGF-1, ER alpha and beta and PR genes, the COX2 gene related to inflammation, GPR54 which is related to the regulation of gonadotropins affecting onset of puberty(34), and CYP3A4*1B which plays a major role in testosterone metabolism and the high activity allele (i.e. CYP3A4*1B) may cause a larger drop in testosterone which may then increase the estradiol: testosterone ratio initiating the hormonal cascade that accompanies puberty(35).

The genes and number of htSNPs for each one that were successfully assayed during no cost extension are listed below (Total=368).

Genes	Number htSNPs
In Original proposal:	
CYP17	11
CYP19	55
COMT	28
HSD17B1	10
CYP1A1	5
GSTP1	8
Additional genes:	
AIB1	27
P160	15
IGF-1	28
IGFBP-3	15
CYP3A4	10
GPR54	10
ESRalpha	83
ESRbeta	21
PR	31
COX-2	11
Total	368

Dr. Dubeau's laboratory did not have the capacity to complete this work and so the work was done under the direction of Dr. David Vandenberg in the Genomics Core Facility. DNA was re-extracted from available samples (i.e. archived tissue or buccal smears) and assays were run for the 368 SNPs using the Illumina System as described below. In addition DNA has been stored from these sample for future testing.

Illumina System Methodology

GoldenGate™ Assay and BeadArray™ Technology

Identification of multiple SNPs at the same time is performed using the GoldenGate™ Assay (Illumina, San Diego, CA). The assay utilizes a combination of the multiplexed oligonucleotide ligation assay (OLA) on genomic DNA (gDNA) and PCR amplification with universal primers. For each polymorphism, two allele specific oligonucleotides (ASO) are synthesized that contain 2 sequence motifs: common sequences at the 5' end for amplification of all targets (P1 and P2) and sequences at the 3' end that match the locus adjacent to the polymorphism with the final base of each oligonucleotide incorporating one of the 2 polymorphic bases. In addition to the 2 allele specific oligonucleotides a locus specific oligonucleotide (LSO) is synthesized that contains 3 sequence motifs: at the 5'end is sequence adjacent to the SNP being evaluated, a locus specific region in the middle of the oligonucleotide

to identify the locus (Address), and sequences at the 3' end for amplification of all targets (P3). During the OLA each allele specific oligonucleotide will anneal to the region next to the corresponding polymorphism and each locus specific oligonucleotide will anneal to the adjacent region downstream of the polymorphism. When the last base of each ASO matches the polymorphic base DNA ligase will ligate the ASO and LSO oligonucleotides together. If a mismatch occurs the ligation step will not occur. Since each locus is independent, a large number of simultaneous annealings can occur provided there is no interaction between the combined oligonucleotides. At present combinations of up to 1536 loci can be performed at once. Next, the ligated oligonucleotides are amplified using generic primers that recognize the common domains within the ASO and LSO oligonucleotides. A total of 3 primers are used to amplify all of the loci at once: 2 primers that are labeled with distinct fluorochromes and are complementary to the P1 and P2 regions, respectively, for each ASO and 1 primer that is complementary to the P3 region of the LSO. Following PCR amplification of the ligation products, the products are denatured and hybridized to an array containing oligonucleotides with sequences complementary to the addresses used to mark each locus in the multiplexing reaction. The array contains approximately 50,000 independent sites with each of the addresses being represented at least 8 times. The array is then read to determine the fluorescent signal present at each address (BeadArray Reader, Illumina). The current system uses a 96-well plate format to detect the genotyping reactions for up to 1,536 assays at a time or 147,456 genotypes per plate. The robotics platform dedicated to the Illumina system is capable of processing at least 6 96-well plates per day for a throughput of over 800,000 genotypes per day. Data from the BeadArray Reader is downloaded to a Laboratory Information Management System (LIMS) and the genotypes are determined using Autogenopipe (Illumina). Genotyped data is retrieved from the LIMS database for analysis.

Assay Design

SNP design will be performed by Illumina from a list of SNPs provided to them for this project. The assay conversion rate for development of a successful assay from an identified SNP is approximately 97% when multiplexing 1,152 SNPs at a time and using "double-hit" SNPs (Fan et al., 2003). Assuming a similar assay conversion rate for this study of known functional SNPs and HapMap identified SNPs we would expect 366 SNPs to work on the Illumina platform (97% of 378). Any SNPs that fail the Illumina design process will be analyzed using the TaqMan assay.

Quality Control

The Genomics Core Facility incorporates 2 levels of Quality control into all assays. Within the sample set a 5-10% blinded duplication of samples is created. Samples will be split and separate IDs generated prior to submitting the samples to the Genomics Core Facility. Results for an assay will not be analyzed if the duplicates do not have identical genotype and the cause for the discordancy (systematic or isolated) will be determined. A second level of QC is provided during sample setup. All DNA samples are diluted and stored in 96-well plates prior to aliquoting of DNA into assay plates. Only 93 samples are added to each 96-well plate with the remaining 3 empty wells serving as negative controls for the assay and as a unique fingerprint

for each 96-well plate. These unique fingerprint wells allow the Genomics Core Facility to identify plate flips, or errors in the creation of assay plates.

References

Jian-Bing Fan, Arnold Oliphant, Richard Shen, Bahram G. Kermani, Francisco Garcia, Kevin L. Gunderson, Mark Hansen, Frank Steemers, Scott L. Butler, Panos Deloukas, Luana Galver, Sarah Hunt, Celeste McBride, Marina Bibikova, Todd Rubano, Jing Chen, Eliza Wickham, Dennis Doucet, Weihua Chang, Derek Campbell, Baohong Zhang, Semyon Kruglyak, David Bentley, Juergen Haas, Philippe Rigault, Lixin Zhou, John Stuelpnagel and Mark S. Chee. Highly Parallel SNP Genotyping. Cold Spring Harbor Symposia on Quantitative Biology, Volume LXVIII, 69-78, January 2004 © 2003 Cold Spring Harbor Laboratory Press.

Task 5 Data analysis

1. *Link data on genetic factors to other information from twins and controls including risk factor information and other tumor related information when available (e.g. ER positivity)*
2. *Complete analyses of data to determine relationship of the specified polymorphisms to breast cancer susceptibility.*
3. *Submit papers and reports.*

Due to poor quality DNA from the archived tissue blocks from some of the twins, all of the samples were not useable. For individual SNPs there were some samples that gave uninformative results for some SNPs but not for others. For each of the 368 SNPs tested we obtained useable results from between 48-71 concordant pairs, 68-99 discordant pairs and 119-127 control women. Frequency distributions of the genotypes and alleles were provided for each group. The chi-square statistic was used to determine if the distributions between concordant and discordant pairs were significantly different and also if there was lack of independence among concordant pairs, discordant pairs and control women.

The number of SNPs that had significantly different distributions are shown in Table 1 and the actual p values, coordinates and rs numbers for each SNP are shown in Table 2. The coordinates indicate the position of the SNP along the gene, thus there may be significance in SNPs that are located close to each other. The comparison between concordant and discordant pairs only showed significance for less than 5% of the SNPs. Since multiple comparisons are being made this could be expected by chance. More SNPs (15-17%) showed significant results when comparing all three strata—concordant pairs, discordant pairs and controls. The only gene where no significant results were found was the GPR54 gene. Genes that had the highest proportion (i.e. >20%) of SNPs that differed in their genotype and allele distributions included HSD17B1, GSTP1 and P160 for the comparison of concordant and discordant pairs, and HSD17B1, CYP1A1, GSTP1, AIB1, and COX2 for the 3-way comparison of concordant and discordant pairs and controls. At this point in time, the meaning of the significance of individual SNPs has not been evaluated, and much more extensive analysis of haplotypes for each gene will need to be done. These results are thus very preliminary at this time.

Comparison of replicates within the samples showed that there was a 1-2% error between samples from the same individual which is acceptable. The Hardy-Weinberg distribution of genotypes in the controls will need to be assessed. The functional significance of the significant SNPs will need to be assessed. In addition, results from CEPH individuals were also included in the assays and a comparison of their results to the known standards will need to be assessed. Further study will also need to be made on combinations of significant SNPs along biologic pathways.

Table 1: Number of SNPs with significant differences in the distribution of genotypes or alleles and the percentage of the total number of SNPs for that gene.

Genes	Total number htSNPs	Number significant (p<.05)			
		Concordant vs. Discordant Pairs		Conc., Disc, and Controls	
		Genotype	Allele	Genotype	Allele
CYP17	11	0	1 (9.1%)	2 (18.2%)	1 (9.1%)
CYP19	55	1 (1.8%)	2 (3.6%)	8 (14.6%)	9 (16.4%)
COMT	28	1 (3.6%)	1 (3.6%)	4 (14.3%)	6 (21.4%)
HSD17B1	10	3 (30.0%)	2 (20.0%)	2 (20.0%)	1 (10.0%)
CYP1A1	5	0 (0%)	0 (0%)	1 (20.0%)	1 (20.0%)
GSTP1	8	1(12.5%)	2 (25.0%)	2 (25.0%)	2 (25.0%)
AIB1	27	1 (3.7%)	2 (7.4%)	6 (22.2%)	6 (22.2%)
P160	15	1 (6.7%)	2 (13.3%)	1 (6.7%)	2 (13.3%)
IGF-1	28	0 (0%)	0 (0%)	3 (10.7%)	5 (17.9%)
IGFBP-3	15	0 (0%)	0 (0%)	1 (6.7%)	2 (13.3%)
CYP3A4	10	0 (0%)	0 (0%)	4 (40.0%)	4 (40.0%)
GPR54	10	0 (0%)	0 (0%)	0 (0%)	0 (0%)
ESRalpha	83	4 (4.8%)	5 (6.0%)	10 (12.0%)	15 (18.1%)
ESRbeta	21	2 (9.5%)	0 (0%)	3 (14.3%)	4 (19.0%)
PR	31	0 (0%)	0 (0%)	3 (9.7%)	2 (6.4%)
COX-2	11	0 (0%)	0 (0%)	4 (36.4%)	3 (27.3%)
Total	368	14 (3.8%)	17 (4.6%)	54 (14.7%)	63 (17.1%)

Table 2: Significance of difference in distributions of genotypes and alleles of htSNPs from 16 genes: Concordant vs. Discordant Pairs and Concordant pairs, Discordant pairs, and Controls. (p values are from Chi square test).

				Concordant vs. Discordant Pairs		Concordant & Discordant Pairs and Controls	
GENE	SNPNO (in database)	COORD	rs number	P value for genotype distribution.	P value for allele distribution	P value for genotype distribution	P value for allele distribution
AIB1	183	45554763	rs2868804	0.81	0.6	0.71	0.8
AIB1	106	45556766	rs17790738	0.95	0.96	0.97	0.86
AIB1	1	45558673	rs13043637	0.83	0.77	0.8	0.9
AIB1	288	45560736	rs6018511	0.1	0.07	0.14	0.1
AIB1	51	45580973	rs1206882	0.54	0.68	0.56	0.92
AIB1	46	45586555	**rs11700063**	0.66	0.84	**0.012**	**0.0193**
AIB1	150	45590796	**rs2425941**	0.37	0.69	**0.0154**	0.34
AIB1	291	45619734	rs6125042	0.38	0.14	0.34	0.09
AIB1	89	45619978	rs1569438	0.45	0.34	0.48	0.38
AIB1	151	45642909	rs2425975	0.37	0.09	0.14	0.09
AIB1	152	45653334	rs2425977	0.25	0.12	0.39	0.23
AIB1	124	45662074	rs2143491	0.5	0.33	0.16	0.11
AIB1	289	45691984	rs6018600	0.27	0.38	0.49	0.49
AIB1	129	45698295	rs2230782	0.17	0.84	0.3	0.98
AIB1	248	45701357	rs4810648	0.7	0.41	0.89	0.62
AIB1	120	45701900	**rs2076546**	0.07	0.09	**0.0277**	**0.0377**
AIB1	290	45708496	**rs6018617**	**0.0443**	**0.0108**	**0.0183**	**0.0027**
AIB1	352	45708735	rs864338	0.34	0.53	0.31	0.19
AIB1	45	45716790	**rs11699879**	0.34	0.16	**0.0159**	**0.0314**
AIB1	232	45719646	**rs445219**	0.2	**0.0176**	0.26	**0.0461**
AIB1	141	45721973	rs2294891	0.54	0.45	0.76	0.6
AIB1	138	45723657	rs2281279	0.75	0.6	0.52	0.7
AIB1	121	45723723	rs2076549	0.72	0.8	0.52	0.81
AIB1	234	45724889	rs450110	0.5	0.8	0.79	0.84
AIB1	131	45725556	rs2235734	0.27	0.78	0.63	0.96
AIB1	119	45728132	rs2076545	0.35	0.18	0.33	0.33
AIB1	228	45729011	**rs403321**	0.17	0.61	**0.0001**	**0.0004**
COMT	347	18283980	rs8141691	0.44	0.21	0.5	0.42
COMT	8	18284531	**rs1012157**	**0.0017**	0.65	**0.0024**	0.87
COMT	326	18285270	rs7289747	0.61	0.41	0.87	0.64
COMT	360	18286569	**rs9306229**	0.2	0.55	**0.0216**	**0.0573**
COMT	285	18289880	rs5993875	0.34	0.51	0.16	0.57
COMT	327	18290493	rs7290448	0.13	0.22	0.3	0.4
COMT	230	18292410	rs4333017	0.89	0.63	0.94	0.66
COMT	233	18293959	**rs4485648**	0.08	**0.0221**	0.07	**0.0468**
COMT	379	18295691	rs9605030	0.12	0.25	0.14	0.17
COMT	117	18303438	**rs2020917**	0.28	0.27	0.09	**0.0328**
COMT	328	18304663	rs737866	0.18	0.31	0.42	0.6
COMT	88	18306222	rs1544325	0.25	0.21	0.46	0.36
COMT	99	18308605	rs174675	0.08	0.0557	0.13	0.16

GENE	SNPNO (in database)	COORD	rs number	Concordant vs. Discordant Pairs		Concordant & Discordant Pairs and Controls	
				P value for genotype distribution.	P value for allele distribution	P value for genotype distribution	P value for allele distribution
COMT	286	18312192	rs5993883	0.97	0.98	0.69	0.72
COMT	329	18319731	rs740603	0.39	0.54	0.47	0.78
COMT	132	18324982	rs2239393	0.7	0.51	0.86	0.8
COMT	242	18325825	rs4680	0.5	0.68	0.76	0.68
COMT	237	18326686	**rs4646316**	0.57	0.35	**0.0122**	**0.0017**
COMT	91	18327115	rs165774	0.86	0.9	0.93	0.93
COMT	100	18327730	rs174696	0.46	0.36	0.8	0.64
COMT	363	18330246	rs9332377	0.52	0.26	0.36	0.43
COMT	92	18333223	rs165849	0.59	0.52	0.61	0.7
COMT	287	18334300	**rs5993891**	0.28	0.34	**0.0001**	**0.0001**
COMT	171	18334742	rs2518823	0.27	1	0.07	0.26
COMT	134	18335564	rs2240714	0.92	0.67	0.99	0.88
COMT	355	18336509	rs887199	0.7	0.73	0.36	0.72
COMT	133	18336757	rs2239395	0.87	0.87	0.98	0.98
COMT	356	18338220	**rs887200**	0.31	0.13	0.15	**0.0285**
COX2	31	183363974	**rs10911902**	0.3	0.56	**0.0044**	**0.0402**
COX2	241	183380661	rs4648261	0.56	0.56	0.81	0.81
COX2	310	183382408	**rs689466**	0.11	0.31	**0.0003**	**0.0146**
COX2	49	183383533	rs12042763	0.46	0.8	0.69	0.96
COX2	178	183383659	**rs2745559**	0.26	0.67	**0.0477**	0.85
COX2	32	183384052	rs10911905	0.71	0.53	0.9	0.7
COX2	57	183388928	rs12409744	0.4	0.33	0.73	0.57
COX2	36	183389651	rs1119064	0.67	0.61	0.7	0.69
COX2	37	183389729	rs1119065	0.17	0.08	0.46	0.2
COX2	307	183391516	**rs6681231**	0.36	0.09	**0.0008**	**0.0025**
COX2	147	183400749	rs2383529	0.6	0.35	0.77	0.64
CYP17	295	104571278	rs619824	0.52	0.99	0.42	0.76
CYP17	30	104573922	rs10883782	0.3	0.12	0.14	0.054
CYP17	253	104574320	rs4919682	0.46	0.88	0.3	0.76
CYP17	97	104581383	rs17115100	0.31	0.88	0.67	0.99
CYP17	293	104586914	rs6163	0.69	0.82	0.46	0.91
CYP17	170	104587470	rs2486758	0.73	0.8	0.89	0.74
CYP17	103	104595511	**rs17724534**	0.1	**0.021**	**0.0251**	0.07
CYP17	315	104599666	rs7096475	0.29	0.75	0.37	0.81
CYP17	55	104603345	rs12219246	0.89	0.62	0.09	0.87
CYP17	224	104604340	**rs3824754**	0.98	0.99	**0.0001**	**0.0001**
CYP17	254	104606490	rs4919690	0.33	0.35	0.33	0.6
CYP19	384	49279146	rs9972359	0.42	0.66	0.74	0.86
CYP19	369	49283122	**rs934632**	0.45	0.74	**0.0401**	**0.0532**
CYP19	370	49287786	rs934633	0.42	0.55	0.79	0.82
CYP19	229	49288409	**rs4275794**	0.72	0.83	**0.0019**	**0.0058**

				Concordant vs. Discordant Pairs		Concordant & Discordant Pairs and Controls	
GENE	SNPNO (in database)	COORD	rs number	P value for genotype distribution.	P value for allele distribution	P value for genotype distribution	P value for allele distribution
CYP19	185	49290969	rs2899470	0.72	0.75	0.6	0.84
CYP19	140	49294800	rs2289105	0.5	0.65	0.14	0.06
CYP19	184	49295260	rs28757190	0.42	0.33	0.2	0.17
CYP19	142	49295412	rs2304463	0.72	0.86	0.25	0.89
CYP19	42	49297994	**rs1143704**	0.44	0.35	**0.01**	**0.0185**
CYP19	186	49303347	rs2899472	0.48	0.75	0.44	0.89
CYP19	344	49305662	rs8025374	0.81	0.59	0.51	0.24
CYP19	78	49310451	rs12900487	0.63	0.92	0.13	0.11
CYP19	69	49312465	rs12592697	0.65	0.33	0.13	0.24
CYP19	383	49315372	rs9944225	0.47	0.93	0.31	0.98
CYP19	314	49316404	rs700518	0.35	0.8	0.32	0.3
CYP19	148	49317127	**rs2414097**	0.84	0.97	**0.0371**	0.13
CYP19	102	49317389	rs17703883	0.86	0.92	0.12	0.14
CYP19	23	49319939	**rs10519295**	0.34	0.64	0.07	**0.049**
CYP19	247	49323314	rs4775936	0.56	0.64	0.07	0.16
CYP19	20	49323433	rs10459592	0.21	0.68	0.45	0.89
CYP19	68	49326660	rs12591359	0.36	0.16	0.53	0.34
CYP19	79	49330049	rs12911554	0.78	0.81	0.32	0.23
CYP19	50	49332163	rs12050772	0.55	0.59	0.52	0.59
CYP19	322	49333590	rs7172156	0.93	0.73	0.83	0.9
CYP19	48	49335997	rs11856927	0.8	0.66	0.75	0.85
CYP19	149	49336074	rs2414099	0.91	0.81	0.81	0.82
CYP19	235	49336336	rs4545755	0.68	0.39	0.76	0.56
CYP19	24	49338638	rs10519299	0.98	0.92	0.58	0.54
CYP19	101	49341201	rs17601876	0.61	0.3	0.24	0.56
CYP19	236	49343886	rs4614671	0.32	0.39	0.21	0.19
CYP19	231	49344251	rs4441215	0.74	0.56	0.93	0.76
CYP19	60	49349490	rs12437685	0.85	0.73	0.83	0.88
CYP19	116	49358145	**rs1902586**	0.11	**0.0311**	0.23	0.09
CYP19	321	49365886	rs7167343	0.31	0.31	0.4	0.41
CYP19	371	49366890	**rs936306**	0.45	0.36	**0.0002**	**0.0008**
CYP19	163	49377192	rs2470157	0.63	0.65	0.69	0.49
CYP19	162	49382264	rs2470152	0.61	0.94	0.74	0.98
CYP19	218	49393870	**rs3751592**	0.22	0.38	**0.0001**	**0.0033**
CYP19	217	49394002	rs3751591	0.85	0.99	0.76	0.65
CYP19	7	49400944	rs1004983	0.64	0.67	0.92	0.91
CYP19	115	49401198	rs1902585	0.98	0.88	0.96	0.76
CYP19	153	49402908	rs2445761	0.29	0.14	0.5	0.33
CYP19	154	49405000	rs2445762	0.85	0.64	0.21	0.08
CYP19	161	49409017	rs2470144	0.8	0.54	0.3	0.32
CYP19	323	49409420	rs7174997	0.94	0.96	0.56	0.94
CYP19	343	49411077	rs8025191	0.61	0.43	0.49	0.25
CYP19	110	49412515	rs1870049	0.26	0.09	0.24	0.059
CYP19	155	49422190	rs2445765	0.11	0.7	0.15	0.9

GENE	SNPNO (in database)	COORD	rs number	Concordant vs. Discordant Pairs		Concordant & Discordant Pairs and Controls	
				P value for genotype distribution.	P value for allele distribution	P value for genotype distribution	P value for allele distribution
CYP19	156	49427651	rs2445771	0.33	0.13	0.7	0.33
CYP19	158	49428448	rs2446426	0.31	0.12	0.39	0.28
CYP19	61	49428833	**rs12441382**	0.55	0.37	**0.0001**	**0.0001**
CYP19	320	49432067	rs7163193	0.17	0.07	0.18	0.09
CYP19	157	49434085	**rs2446405**	**0.0316**	**0.0066**	**0.0375**	**0.0115**
CYP19	123	49435826	**rs2124873**	0.29	0.16	0.052	**0.0082**
CYP19	2	49437643	rs2470184	0.37	0.29	0.12	0.08
CYP1A1	238	72803245	rs4646421	0.38	0.41	0.68	0.71
CYP1A1	164	72806502	rs2470893	0.7	0.91	0.59	0.44
CYP1A1	167	72814933	**rs2472297**	0.74	0.48	**0.0001**	**0.0001**
CYP1A1	95	72819640	rs16972208	0.43	0.43	0.73	0.74
CYP1A1	168	72820453	rs2472299	0.64	0.66	0.49	0.87
CYP3A4	56	98998765	**rs12333983**	0.92	0.66	**0.0119**	**0.0056**
CYP3A4	175	99001488	rs2687126	0.74	0.79	**0.0001**	**0.0005**
CYP3A4	135	99006117	rs2242480	0.32	0.54	0.65	0.72
CYP3A4	239	99009734	rs4646437	0.97	0.98	0.81	0.48
CYP3A4	176	99015144	**rs2738258**	0.46	0.96	**0.0001**	**0.0001**
CYP3A4	13	99019161	rs10270146	0.76	1	0.27	1
CYP3A4	177	99026747	rs2740574	0.43	0.19	0.67	34
CYP3A4	109	99027587	rs1851426	0.74	0.41	0.69	0.67
CYP3A4	330	99034426	**rs760368**	0.88	0.89	**0.0001**	**0.0001**
CYP3A4	126	99040725	rs2177179	0.36	0.21	0.16	0.1
ESR1	270	152203057	rs543650	0.24	0.54	0.5	0.8
ESR1	375	152209572	rs9478243	0.45	0.94	0.8	0.95
ESR1	376	152214151	rs9478244	0.76	0.59	0.32	0.53
ESR1	252	152217558	rs488133	0.36	0.97	0.57	0.79
ESR1	268	152223032	rs532010	0.63	0.82	0.5	0.33
ESR1	22	152224431	rs10484922	0.7	0.59	0.39	0.71
ESR1	225	152232000	**rs3853248**	0.8	0.88	**0.0015**	**0.0005**
ESR1	73	152241986	rs12665044	0.55	0.29	0.1	0.054
ESR1	335	152243977	rs7761133	0.27	0.12	0.12	0.3
ESR1	350	152248311	rs827423	0.66	0.44	0.39	0.16
ESR1	226	152252014	rs3853250	0.76	0.73	0.75	0.74
ESR1	361	152254431	rs9322331	0.83	0.5	0.2	0.08
ESR1	130	152255449	rs2234693	0.58	0.72	0.09	0.14
ESR1	227	152259425	**rs3936674**	0.36	0.52	0.06	**0.0213**
ESR1	349	152269643	**rs827420**	0.4	0.67	0.06	**0.0303**
ESR1	348	152269777	rs827419	0.63	0.9	0.86	0.89
ESR1	96	152286110	rs1709183	0.78	0.46	0.97	0.74
ESR1	35	152291473	rs11155819	0.47	0.36	0.65	0.46
ESR1	362	152292544	**rs9322336**	0.06	**0.0331**	0.18	0.06

GENE	SNPNO (in database)	COORD	rs number	Concordant vs. Discordant Pairs		Concordant & Discordant Pairs and Controls	
				P value for genotype distribution.	P value for allele distribution	P value for genotype distribution	P value for allele distribution
ESR1	336	152304622	rs7761846	0.22	0.17	0.21	0.16
ESR1	300	152326707	**rs6557171**	**0.0431**	**0.0103**	**0.0107**	**0.0139**
ESR1	250	152329582	rs4870061	0.2	0.12	0.29	0.14
ESR1	251	152329732	**rs4870062**	**0.0379**	**0.0345**	**0.0022**	**0.0073**
ESR1	381	152333264	rs988328	0.41	0.16	0.41	0.19
ESR1	373	152339266	**rs9397456**	0.57	0.36	0.16	**0.0371**
ESR1	312	152349801	**rs6927072**	0.11	0.57	**0.0066**	0.74
ESR1	372	152354227	**rs9371564**	0.08	0.057	**0.0305**	**0.0142**
ESR1	108	152357636	rs1801132	0.48	0.37	0.69	0.38
ESR1	208	152358491	rs3020410	0.07	0.94	0.07	0.97
ESR1	198	152358582	rs3003917	0.35	0.16	0.35	0.14
ESR1	311	152360654	rs6914211	0.14	0.15	0.37	0.3
ESR1	203	152364512	rs3020377	0.61	0.85	0.9	0.98
ESR1	200	152370855	rs3020317	0.63	0.33	0.29	0.22
ESR1	205	152375158	rs3020401	0.76	0.92	0.94	0.96
ESR1	111	152375393	rs1884051	0.53	0.56	0.8	0.8
ESR1	5	152375592	rs985192	0.84	0.9	0.96	0.99
ESR1	199	152376572	**rs3003925**	0.31	0.18	0.06	**0.0277**
ESR1	301	152376935	rs6557177	0.91	0.73	0.99	0.91
ESR1	192	152378637	rs2982700	0.89	0.54	0.6	0.82
ESR1	380	152378739	rs985694	0.67	0.42	0.3	0.31
ESR1	201	152381884	rs3020318	0.92	0.82	0.95	0.92
ESR1	112	152383480	rs1884052	0.9	0.93	0.24	0.99
ESR1	113	152383680	rs1884054	0.36	0.61	0.59	0.69
ESR1	206	152387829	rs3020403	0.74	0.88	0.96	0.99
ESR1	191	152390549	rs2982683	0.77	0.91	0.69	0.6
ESR1	374	152396442	**rs9397463**	0.6	0.45	**0.0333**	**0.0264**
ESR1	359	152397161	rs926777	0.92	0.95	0.96	0.99
ESR1	202	152397619	rs3020328	0.35	0.12	0.57	0.27
ESR1	207	152399375	rs3020407	0.81	0.73	0.87	0.92
ESR1	125	152399820	**rs2144025**	0.06	**0.0175**	0.1	**0.0218**
ESR1	331	152401246	rs7743290	0.38	0.45	0.65	0.75
ESR1	54	152402121	rs12212176	0.95	0.77	0.34	0.24
ESR1	332	152403651	rs7754762	0.62	0.47	0.74	0.69
ESR1	364	152405260	rs9340941	0.51	0.21	0.38	0.38
ESR1	334	152409254	rs7757956	0.1	0.89	0.09	0.5
ESR1	365	152412286	rs9340954	0.18	0.16	0.2	0.32
ESR1	90	152420730	rs1569788	0.19	0.37	0.24	0.42
ESR1	366	152422315	rs9340955	0.11	1	0.08	1
ESR1	81	152425218	rs13203975	0.79	0.5	0.92	0.74
ESR1	333	152431729	**rs7755185**	0.32	0.09	**0.0352**	0.17
ESR1	209	152435877	rs3020411	0.96	0.87	0.98	0.91
ESR1	3	152448334	rs2982708	0.63	0.95	0.56	0.71
ESR1	193	152448858	rs2982709	0.35	0.28	0.21	0.0551

GENE	SNPNO (in database)	COORD	rs number	Concordant vs. Discordant Pairs		Concordant & Discordant Pairs and Controls	
				P value for genotype distribution.	P value for allele distribution	P value for genotype distribution	P value for allele distribution
ESR1	210	152450041	rs3020432	0.66	0.44	0.42	0.64
ESR1	194	152450293	rs2982712	0.3	0.11	0.47	0.2
ESR1	211	152451054	rs3020434	0.28	0.24	0.13	0.2
ESR1	367	152474802	rs9341019	0.5	0.25	0.56	0.25
ESR1	195	152484156	rs2982894	0.95	0.7	0.41	0.38
ESR1	53	152484712	rs12199198	0.52	0.23	0.68	0.38
ESR1	52	152486893	rs12180788	0.7	0.37	0.87	0.52
ESR1	196	152491607	**rs2982896**	0.91	0.83	0.11	**0.0178**
ESR1	169	152505018	rs2474148	0.77	0.83	0.17	0.58
ESR1	197	152507106	**rs2982900**	**0.0108**	0.33	**0.0012**	**0.0207**
ESR1	368	152508739	**rs9341052**	**0.0087**	**0.01**	**0.0376**	**0.0425**
ESR1	204	152508893	rs3020383	0.56	0.71	0.14	**0.0044**
ESR1	220	152510689	rs3778099	0.98	0.98	0.84	0.54
ESR1	128	152512209	rs2228480	0.97	0.98	0.9	0.89
ESR1	221	152513244	rs3798577	0.2	0.12	0.16	0.17
ESR1	180	152516592	**rs2813543**	0.67	0.37	**0.0272**	**0.0054**
ESR1	181	152517696	rs2813544	0.99	0.95	0.98	0.99
ESR1	179	152518615	rs2747649	0.89	0.97	0.47	0.97
ESR1	87	152520818	rs1543403	0.46	0.39	0.09	0.14
ESR1	358	152525016	rs910416	0.75	0.74	0.49	0.23
ESR2	58	63761606	rs12434245	0.5	0.74	0.49	0.43
ESR2	44	63762383	rs1152582	0.28	0.46	0.17	0.17
ESR2	63	63763624	rs1255998	0.54	0.95	0.66	0.99
ESR2	342	63763835	rs8018687	0.91	0.83	0.9	0.84
ESR2	341	63769203	rs8006145	0.96	0.93	0.62	0.98
ESR2	259	63769569	rs4986938	0.16	0.73	0.18	0.48
ESR2	354	63770795	rs867443	0.67	0.78	0.76	0.75
ESR2	67	63771970	**rs1256063**	0.94	0.88	0.055	**0.0428**
ESR2	66	63773071	rs1256062	0.33	0.83	0.42	0.94
ESR2	65	63780170	rs1256059	0.56	0.78	0.65	0.66
ESR2	137	63786382	rs2274705	0.64	0.97	0.42	0.3
ESR2	59	63793278	**rs12435857**	**0.016**	0.32	**0.0069**	**0.0422**
ESR2	340	63795122	rs8003490	0.96	0.76	0.67	0.46
ESR2	64	63803780	**rs1256044**	0.22	0.8	0.052	**0.0455**
ESR2	318	63806413	**rs7154455**	0.97	0.99	0.2	**0.0189**
ESR2	378	63814932	rs960070	0.82	0.96	0.06	0.14
ESR2	98	63826504	**rs17179740**	**0.0331**	0.99	**0.0128**	0.08
ESR2	319	63828629	rs7159462	0.41	0.28	0.33	0.19
ESR2	114	63830364	rs1887994	0.92	0.98	0.99	0.99
ESR2	76	63831670	**rs1271572**	0.42	0.81	**0.0233**	0.28
ESR2	9	63845529	rs10137185	0.93	0.95	0.15	0.06
GPR54	216	848571	rs3746149	0.14	0.19	0.21	0.39

GENE	SNPNO (in database)	COORD	rs number	Concordant vs. Discordant Pairs		Concordant & Discordant Pairs and Controls	
				P value for genotype distribution.	P value for allele distribution	P value for genotype distribution	P value for allele distribution
GPR54	143	850978	rs2306718	0.7	0.37	0.1	0.66
GPR54	346	857633	rs8112519	0.17	0.052	0.33	0.09
GPR54	16	858673	rs10425660	0.99	0.99	0.45	0.58
GPR54	93	861742	rs168405	0.47	0.78	0.21	0.27
GPR54	213	874744	rs350134	0.43	0.37	0.69	0.48
GPR54	345	878971	rs8108687	0.69	0.43	0.67	0.71
GSTP1	309	67096525	rs688878	0.38	0.16	0.17	0.16
GSTP1	292	67103863	rs614080	0.17	0.0585	0.38	0.12
GSTP1	338	67104171	**rs7941648**	**0.025**	**0.0095**	**0.0164**	**0.0056**
GSTP1	302	67106475	**rs6591256**	0.1	**0.0319**	0.22	0.09
GSTP1	377	67109265	rs947894	0.17	0.09	0.1	0.11
GSTP1	107	67110155	**rs1799811**	0.62	0.54	**0.0015**	**0.001**
GSTP1	283	67116179	rs596603	0.06	0.27	0.19	0.51
GSTP1	41	67116840	rs11227844	0.87	1	0.35	0.99
HSD17B1	303	37942981	**rs659497**	**0.018**	**0.0193**	**0.0034**	**0.0037**
HSD17B1	118	37943139	**rs2071046**	**0.0068**	0.74	**0.0078**	0.94
HSD17B1	299	37946870	**rs630539**	0.12	**0.024**	0.41	0.08
HSD17B1	351	37949759	rs86312	0.07	0.97	0.17	0.99
HSD17B1	182	37958089	rs2830	0.58	0.86	0.22	0.21
HSD17B1	284	37958626	rs597255	0.32	0.92	0.45	0.67
HSD17B1	172	37959481	rs2676530	0.81	0.65	0.9	0.9
HSD17B1	308	37959799	rs676387	0.81	0.92	0.89	0.98
HSD17B1	70	37965295	rs12602084	0.31	0.5	0.35	0.2
HSD17B1	296	37965943	**rs621141**	**0.0473**	1	0.26	0.27
IGF1	144	101282429	rs2373720	0.99	0.99	0.11	0.13
IGF1	190	101283346	rs2971575	0.82	0.53	0.61	0.66
IGF1	28	101288036	**rs10860861**	0.25	0.12	**0.0159**	**0.0078**
IGF1	29	101288539	rs10860862	0.78	0.91	0.34	0.99
IGF1	188	101290281	rs2946834	0.95	0.83	0.69	0.51
IGF1	298	101292659	rs6219	0.12	0.73	0.17	0.82
IGF1	297	101296036	**rs6214**	0.7	0.5	**0.0252**	**0.0034**
IGF1	86	101298989	rs1520220	0.4	0.98	0.35	0.99
IGF1	104	101312097	rs17727841	0.89	0.92	0.33	0.99
IGF1	261	101312736	rs5009837	0.18	0.28	0.48	0.54
IGF1	278	101314993	**rs5742688**	0.36	0.16	0.19	**0.0454**
IGF1	145	101329512	rs2373721	0.8	0.99	0.82	0.99
IGF1	139	101332475	rs2288378	0.84	0.65	0.55	0.28
IGF1	277	101338533	rs5742652	0.74	0.74	0.74	0.91
IGF1	317	101340982	**rs7136446**	0.44	0.2	0.15	**0.0405**
IGF1	146	101342924	rs2373722	0.65	0.66	0.07	0.08
IGF1	26	101346703	**rs10735380**	0.94	0.74	0.09	**0.0145**

				Concordant vs. Discordant Pairs		Concordant & Discordant Pairs and Controls	
GENE	SNPNO (in database)	COORD	rs number	P value for genotype distribution.	P value for allele distribution	P value for genotype distribution	P value for allele distribution
IGF1	276	101350713	rs5742639	0.27	0.27	0.51	0.52
IGF1	47	101355710	rs11831436	0.44	0.45	0.76	0.45
IGF1	127	101359169	rs2195239	0.5	0.72	0.44	0.18
IGF1	275	101359730	rs5742629	0.7	0.68	0.14	0.06
IGF1	27	101365446	rs10778176	0.26	0.49	0.11	0.07
IGF1	10	101366892	**rs1019731**	0.91	0.99	**0.004**	0.99
IGF1	6	101370134	rs12821878	0.62	0.77	0.5	0.29
IGF1	274	101372067	rs5742620	0.48	1	0.68	0.99
IGF1	214	101378036	rs35767	0.35	0.38	0.31	0.63
IGF1	382	101382589	rs9919733	0.55	0.56	0.78	0.79
IGF1	353	101396326	rs865927	0.24	0.66	0.3	0.24
IGFBP3	75	45718709	rs12702181	0.94	0.86	0.95	0.89
IGFBP3	74	45720395	rs12671484	0.91	0.69	0.74	0.31
IGFBP3	136	45722810	rs2270628	0.21	0.38	0.41	0.68
IGFBP3	82	45724470	rs13223993	0.99	0.9	0.07	0.19
IGFBP3	306	45725494	rs6670	0.48	0.85	0.71	0.87
IGFBP3	159	45726813	rs2453839	0.86	0.86	0.18	0.08
IGFBP3	12	45727932	**rs10255707**	0.38	0.35	0.2	**0.0396**
IGFBP3	212	45728269	rs3110697	0.15	0.58	0.12	0.16
IGFBP3	357	45738235	rs903889	0.97	93	0.84	0.78
IGFBP3	83	45744350	rs13232606	0.19	0.06	0.07	0.057
IGFBP3	11	45747298	rs10235181	0.65	0.55	0.62	0.68
IGFBP3	160	45747905	**rs2453849**	0.79	0.82	**0.0004**	**0.0104**
IGFBP3	165	45751363	rs2471553	0.11	0.47	0.25	0.52
IGFBP3	166	45752807	rs2471554	0.92	0.81	0.81	0.6
IGFBP3	189	45753991	rs2965072	0.7	0.7	0.77	0.78
P160	84	4380349	**rs1351231**	**0.0449**	0.6	0.09	0.74
P160	80	4383646	rs12949158	0.11	0.97	0.36	0.98
P160	33	4386031	rs11078514	0.64	0.41	0.91	0.68
P160	25	4388058	rs10521140	0.06	0.84	0.11	0.98
P160	19	4388634	rs1045845	0.1	0.32	0.13	0.35
P160	62	4391384	rs12450708	0.07	0.87	0.24	0.98
P160	223	4401915	rs3816686	0.59	0.48	0.61	0.62
P160	71	4405539	rs12603519	0.48	0.39	0.51	0.58
P160	219	4405958	rs3760194	0.34	0.77	0.4	0.95
P160	324	4410448	rs7216284	0.75	0.52	0.08	0.16
P160	325	4410545	rs7216474	0.85	0.91	0.75	0.53
P160	34	4411345	rs11078517	0.44	0.76	0.07	0.2
P160	122	4412251	rs2100986	0.08	0.16	0.08	0.2
P160	15	4421661	**rs1038122**	0.08	**0.0257**	**0.0448**	**0.0086**
P160	14	4421993	**rs1038121**	0.11	**0.0429**	0.14	**0.0413**

GENE	SNPNO (in database)	COORD	rs number	Concordant vs. Discordant Pairs		Concordant & Discordant Pairs and Controls	
				P value for genotype distribution.	P value for allele distribution	P value for genotype distribution	P value for allele distribution
PR	105	100404528	rs17728653	0.27	0.68	0.27	0.76
PR	21	100405926	**rs1046982**	0.32	0.44	**0.0034**	**0.0024**
PR	272	100408204	rs561610	0.53	0.53	0.67	0.6
PR	243	100410507	rs471767	0.57	0.25	0.71	0.48
PR	266	100413084	rs523535	0.91	0.97	0.94	0.77
PR	260	100415201	rs500760	0.86	0.62	0.99	0.88
PR	282	100416758	rs588913	0.97	0.92	0.9	0.87
PR	18	100427412	rs1042839	0.24	0.14	0.57	0.31
PR	279	100427614	rs578029	0.95	0.76	0.97	0.86
PR	38	100429243	rs11224575	0.55	0.76	0.77	0.93
PR	255	100432070	rs492457	0.91	0.8	0.96	0.86
PR	43	100434397	rs1144133	0.61	0.58	0.29	0.82
PR	17	100438622	**rs1042838**	1	1	**0.0081**	0.98
PR	304	100439577	rs660541	0.99	0.93	0.92	0.67
PR	256	100440990	rs495997	0.99	0.96	0.89	0.86
PR	39	100443503	rs11224580	0.71	0.72	0.25	0.11
PR	305	100443654	rs665617	0.18	0.56	0.43	0.61
PR	281	100465785	rs585447	0.58	0.28	0.73	0.54
PR	264	100467410	rs508653	0.32	0.92	0.39	0.99
PR	263	100470449	rs508533	0.54	0.43	0.33	0.23
PR	280	100472265	rs578938	0.6	0.47	0.8	0.77
PR	271	100474755	rs555653	0.17	0.77	0.44	0.96
PR	40	100477150	rs11224589	0.32	0.2	0.56	0.4
PR	294	100477546	rs619487	0.51	0.75	0.13	0.46
PR	85	100487782	rs1456765	0.54	0.9	0.68	0.98
PR	269	100493244	**rs537681**	0.47	0.4	**0.0017**	**0.0062**
PR	249	100495657	rs485283	0.71	0.57	0.86	0.59
PR	265	100505711	rs518162	0.11	0.12	0.21	0.31
PR	262	100508322	rs507141	0.64	0.36	0.91	0.63
PR	246	100513712	rs4754732	0.92	0.73	0.98	0.88
PR	245	100519759	rs474320	0.26	0.13	0.31	0.31

.

6) Key Research Accomplishments

a. We have obtained DNA and signed consent forms for 136 concordant pairs, 152 discordant pairs, and 137 controls.
b. DNA has been extracted from all available samples and stored for future testing.
c. Assays on 368 SNPs along 16 genes have been completed used the Illumina System.
d. Very preliminary assessment of significant differences between the distributions of genotypes and alleles of concordant for breast cancer pairs and discordant for breast cancer pairs has been provided. In addition, significance based on the chi-square statistic has been determined for the distributions of genotypes and alleles for concordant and discordant pairs and control women. Less than 5% of the SNPs showed

significant differences for concordant vs. discordant pairs whereas 15-17% were significant for the three-way comparison of concordant and discordant pairs and controls.

7) Reportable Outcomes

Preliminary results indicate that some of the studied genes may be involved in breast cancer susceptibility, but further analyses are required.

8) Conclusions

We have successfully obtained DNA samples from 136 concordant pairs, 152 discordant pairs and 137 controls for a total of 425 samples. DNA has been extracted and stored for additional genetic testing from these samples. A total of 368 SNPs have been assayed along 16 genes. The genes include AIB1, COMT, COX2, CYP17, CYP19, CYP1A1, CYP3A4, ESR1, ESR2, GPR54 ,GSTP1, IGF1, IGFBP3, P160, and PR. The SNPs selected were essentially haplotype tagging SNPs that were selected to cover the variation across the entire length of each of the genes. The genes that showed the most indication of being involved with breast cancer susceptibility included HSD17B1, CYP1A1, GSTP1, AIB1, P160 and COX2. The project has generated a wealth of data that will require further analysis to understand the significance of these results. This group of twins represents an extremely important and valuable group to study breast cancer susceptibility genes and with the DNA stored as a result study, additional SNPs can be easily tested.

9) References

1. Byrne C, Brinton LA, Haile RW, Schairer C. Heterogeneity of the effect of family history on breast cancer risk. Epidemiology 1991;2:276-84.
2. Castilla LH, Couch FJ, Erdos MR, et al. Mutations in the BRCA1 gene in families with early-onset breast and ovarian cancer. Nature Genetics 1994;8:387-91.
3. Feigelson HS, Ross RK, Yu MC, Coetzee GA, Reichardt JK, Henderson BE. Genetic susceptibility to cancer from exogenous and endogenous exposures. Journal of Cellular Biochemistry - Supplement 1996;25:15-22.
4. Kelsey JL, Gammon MD, John EM. Reproductive factors and breast cancer. Epidemiologic Reviews 1993;15:36-47.
5. Bernstein L, Ross RK. Endogenous hormones and breast cancer risk. Epidemiologic Reviews 1993;15:48-65.
6. Pike MC, Spicer DV, Dahmoush L, Press MF. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. Epidemiologic Reviews 1993;15:17-35.
7. Rebbeck TR. Inherited genetic predisposition in breast cancer. A population-based perspective. Cancer 1999;86:2493-501.
8. Dunning AM, Healey CS, Pharoah PD, Teare MD, Ponder BA, Easton DF. A systematic review of genetic polymorphisms and breast cancer risk. Cancer Epidemiology, Biomarkers & Prevention 1999;8:843-54.
9. Coughlin SS, Piper M. Genetic polymorphisms and risk of breast cancer. Cancer Epidemiology, Biomarkers & Prevention 1999;8:1023-32.
10. Feigelson HS, Coetzee GA, Kolonel LN, Ross RK, Henderson BE. A polymorphism in the CYP17 gene increases the risk of breast cancer. Cancer Research 1997;57:1063-5.
11. Dunning AM, Healey CS, Pharoah PD, et al. No association between a polymorphism in the steroid metabolism gene CYP17 and risk of breast cancer. British Journal of Cancer 1998;77:2045-7.

12. Ford D, Easton DF, Peto J. Estimates of the gene frequency of BRCA1 and its contribution to breast and ovarian cancer incidence. American Journal of Human Genetics 1995;57:1457-62.
13. Rowell S, Newman B, Boyd J, King MC. Inherited predisposition to breast and ovarian cancer [editorial; comment]. American Journal of Human Genetics 1994;55:861-5.
14. Martucci CP, Fishman J. P450 enzymes of estrogen metabolism. Pharmacology & Therapeutics 1993;57:237-57.
15. Feigelson HS, Shames LS, Pike MC, Coetzee GA, Stanczyk FZ, Henderson BE. Cytochrome P450c17alpha gene (CYP17) polymorphism is associated with serum estrogen and progesterone concentrations. Cancer Research 1998;58:585-7.
16. Siegelmann-Danieli N, Buetow KH. Constitutional genetic variation at the human aromatase gene (Cyp19) and breast cancer risk. British Journal of Cancer 1999;79:456-63.
17. Kristensen VN, Andersen TI, Lindblom A, Erikstein B, Magnus P, Borresen-Dale AL. A rare CYP19 (aromatase) variant may increase the risk of breast cancer. Pharmacogenetics 1998;8:43-8.
18. Huang CS, Chern HD, Chang KJ, Cheng CW, Hsu SM, Shen CY. Breast cancer risk associated with genotype polymorphism of the estrogen-metabolizing genes CYP17, CYP1A1, and COMT: a multigenic study on cancer susceptibility. Cancer Research 1999;59:4870-5.
19. Thompson PA, Shields PG, Freudenheim JL, et al. Genetic polymorphisms in catechol-O-methyltransferase, menopausal status, and breast cancer risk. Cancer Research 1998;58:2107-10.
20. Millikan RC, Pittman GS, Tse CK, et al. Catechol-O-methyltransferase and breast cancer risk. Carcinogenesis 1998;19:1943-7.
21. Pasqualini JR. Role, control and expression of estrone sulfatase and 17 beta-hydroxysteroid dehydrogenase activities in human breast cancer. Zentralblatt fur Gynakologie 1997;119:48-53.
22. Miettinen MM, Mustonen MV, Poutanen MH, Isomaa VV, Vihko RK. Human 17 beta-hydroxysteroid dehydrogenase type 1 and type 2 isoenzymes have opposite activities in cultured cells and characteristic cell- and tissue-specific expression. Biochemical Journal 1996;314:839-45.
23. Peltoketo H, Vihko P, Vihko R. Regulation of estrogen action: role of 17 beta-hydroxysteroid dehydrogenases. Vitamins & Hormones 1999;55:353-98.
24. Henderson B, Feigelson HS. Hormonal carcinogenesis. Carcinogenesis 2000;21:427-433.
25. Mannermaa A, Peltoketo H, Winqvist R, et al. Human familial and sporadic breast cancer: analysis of the coding regions of the 17 beta-hydroxysteroid dehydrogenase 2 gene (EDH17B2) using a single-strand conformation polymorphism assay. Human Genetics 1994;93:319-24.
26. Peltoketo H, Nokelainen P, Piao YS, Vihko R, Vihko P. Two 17beta-hydroxysteroid dehydrogenases (17HSDs) of estradiol biosynthesis: 17HSD type 1 and type 7. Journal of Steroid Biochemistry & Molecular Biology 1999;69:431-9.
27. Peltoketo H, Piao Y, Mannermaa A, et al. A point mutation in the putative TATA box, detected in nondiseased individuals and patients with hereditary breast cancer, decreases promoter activity of the 17 beta-hydroxysteroid dehydrogenase type 1 gene 2 (EDH17B2) in vitro. Genomics 1994;23:250-2.
28. Charrier J, Maugard CM, Le Mevel B, Bignon YJ. Allelotype influence at glutathione S-transferase M1 locus on breast cancer susceptibility. British Journal of Cancer 1999;79:346-53.
29. Helzlsouer KJ, Selmin O, Huang HY, et al. Association between glutathione S-transferase M1, P1, and T1 genetic polymorphisms and development of breast cancer [see comments]. Journal of the National Cancer Institute 1998;90:512-8.

30. Huang CS, Shen CY, Chang KJ, Hsu SM, Chern HD. Cytochrome P4501A1 polymorphism as a susceptibility factor for breast cancer in postmenopausal Chinese women in Taiwan. British Journal of Cancer 1999;80:1838-43.
31. Taioli E, Bradlow HL, Garbers SV, et al. Role of estradiol metabolism and CYP1A1 polymorphisms in breast cancer risk. Cancer Detection & Prevention 1999;23:232-7.
32. Deapen D, Escalante A, Weinrib L, et al. A revised estimate of twin concordance in systemic lupus erythematosus [see comments]. Arthritis & Rheumatism 1992;35:311-8.
33. Hamilton AS, Mack TM. Puberty and genetic susceptibility to breast cancer in a case-control study in twins.New England Journal of Medicine. 2003; 348:2313-22.
34. Seminara SB, Messager S, Chatzidaki EE, Thresher RR, Acierno JS, Jr., Shagoury JK, et al: The GPR54 gene as a regulator of puberty. New England Journal of Medicine 1614;349:1614-27.
35. Kadlubar FF, Berkowitz GS, Delongchamp RR, Wang C, Green BL, Tang G, et al: The CYP3A4*1B variant is related to the onset of puberty, a known risk factor for the development of breast cancer. Cancer Epidemiology, Biomarkers & Prevention. 2003;12:327-31
36. Mack TM, Hamilton AS, Press MF, Diep A, Rappaport EB. Heritable breast cancer in twins. *British Journal of Cancer.* 2002;87(3):294-300.
37. Haiman CA, Stram DO, Pike MC, et al. A comprehensive haplotype analysis of CYP19 and breast cancer risk: The Multiethnic Cohort. *Human Molecular Genetics.* 2003;12:2679-2692.